Automatisierte Rechtsvergleichung

Posted on 15. März 2026 by Klaus F. Röhl

Drei neuere Forschungsarbeiten haben meine Aufmerksamkeit auf die Möglichkeiten der Rechtsvergleichung mit Hilfe Künstlicher Intelligenz (KI) gelenkt.

1. *Corinna Coupette/Dirk Hartung*, Rechtsstrukturvergleichung, RabelsZ 2022, 935–975.
2. *Kilian Lüders/Bent Stohlmann*, Classifying Proportionality – Identification of a Legal Argument, Artificial Intelligence and Law 2025, 1051–1078
3. *Corinna Coupette/Dirk Hartung/Janis Beckedorf/Maximilian Böther/Daniel Martin Katz*, Law Smells, Artificial Intelligence Law 31, 2023, 335–368.

Zu 1: *Corinna Coupette/Dirk Hartung, Rechtsstrukturvergleichung,*

Der Name von *Corinna Coupette* erscheint in dieser kurzen Liste doppelt, und eigentlich müsste er noch öfter genannt werden. Mit ihrer erstaunlichen Dissertation über »Juristische Netzwerkforschung« (J. C. B. Mohr, Tübingen, 2019) hat sie eine theoretische Grundlage für die Quantifizierung von Rechtsphänomenen gelegt. Der Titel des Buches bringt nicht zum Ausdruck, dass es um eine mathematische Modellierung als Voraussetzung der automatisierten Erfassung von Rechtsstrukturen geht, kurz um »quantitative Rechtswissenschaft«, nicht aber um den sozialen und soziologischen Aspekt von Netzwerken geht. Was man unter »quantitativer Rechtswissenschaft« zu verstehen hat und was moderne Methoden insoweit leisten können, haben *Corinna Coupette* und *Andreas M. Fleckner* in 2018 in der Juristenzeitung dargestellt. [1] Wie man den Netzwerkansatz nutzen kann, haben *Coupette* u. a. 2021 ausprobiert.[2] Diese empirische Basis nutzen sie, um zu zeigen, wie eine quantifizierende Rechtsvergleichung anhand von (äußeren) Ähnlichkeitsmerkmalen möglich ist. Die Autoren sprechen von »quantitativer Rechtsvergleichung (quantitative comparative law)«.

Quantitative Methoden stehen in den Humanities prinzipiell unter Verdacht, die Dinge unzulässig zu vereinfachen. Das gilt gerade auch für die Rechtsvergleichung. *Coupette* und *Hartung* inszenieren ihre Forschung daher als einen Schritt der Komplexitätsreduktion. Das ist als vorbeugende Verteidigung gegen kulturalistische Angriffe verständlich. Ich würde mir allerdings eine aggressivere Position wünschen.

Dass Rechtssysteme »komplexe adaptive Systeme« bilden, die ein Beobachter in ihrer Gesamtheit nicht nachvollziehen und damit verstehen kann, haben wir längst alle begriffen. Da wir nun aber in solchen Systemen leben und uns in ihnen zurechtfinden müssen, ist es von vornherein ausgeschlossen, stets die ganze Dynamik des Systems in Betracht zu ziehen. Man muss sich zunächst an groben Strukturen orientieren und kann nur bei Bedarf in die Tiefe gehen. Solcher Bedarf ist in erster Linie gegeben, wenn es um das Schicksal bestimmter Menschen geht. Doch selbst dann muss man sich zuvor der »Rahmenbedingungen« versichern.[3]

Der Strukturbegriff ist notorisch unscharf. Er lenkt den Blick auf etwas relativ Festes und Dauerhaftes. Als Elemente der Struktur kann man Rechtsformbegriffe wählen, so dass die Struktur auf einer formalen, äußeren Ebene bleibt. Als Alternative ist eine Struktur aus Werten und Prinzipien denkbar. Empirisch lassen sich Formen leichter beobachten als Inhalte. Das ist der Hintergrund der Arbeit von *Coupette* und *Hartung*. Sie stellen »Objekte« zusammen, die aufgrund ihrer Wechselbeziehungen gegenüber ihrer Umwelt als zusammengehörig erscheinen und so ein System bilden, das sich auch als Netzwerk empirisch und mathematisch erfassen lässt.

Über die Auswahl der »Objekte« wird nicht weiter reflektiert. Die Objekte werden schlicht als »konzeptionell anerkannt« und »funktional fundiert« vorgestellt (S. 952). Die Basis bilden Legislative, Exekutive und Judikative als staatliche Akteure und daneben private, Individuen, Unternehmen) und Literatur(-produzenten). In Betracht kommen Organisationen, Personen und Dokumente, kurz alles, was Gegenstand etwa von Justizstatistiken ist. In Betracht kommen aber auch formgeprägte Differenzierungen wie die zwischen Legislative, Exekutive und Judikative (Gewaltenteilung oder zwischen Gesetz, Verordnung, Vertrag, Verwaltungsakt und Gerichtsurteil (Rechtsquellen). Diese Akteure handeln teils durch Individuen (Abgeordnete, Beamte, Richter, Private, Juristen – das spielt aber in der konkreten Untersuchung keine Rolle), teils in Organisationen (Parlament, Behörde, Gericht, Unternehmen, Forschungsinstitute). Sie produzieren Dokumente, nämlich Gesetze, Verordnungen und Verwaltungsakte, Entscheidungen, Verträge oder juristisches Schrifttum. Schließlich werden die Objekte auf drei Ebenen verortet, auf der Makroebene ganzer Rechtssysteme, auf der mittleren Ebene bestimmter Rechtsgebiete und auf der Mikroebene etwa einzelner Dokumente (Gesetze, Verordnungen usw.). Das ist also die von *Coupette* und *Hartung* verwendete »Taxonomie«, bzw. das sind ihre Variablen.

Die Zusammenhänge zwischen den Objekten lassen sich in Tabellen darstellen, wie es die so genannte Businesss Process Management Notation (BPMN)[4] vorsieht.

Längst gibt es Programme, die BPMN-Tabellen automatisch auswerten können. *Coupette* und *Hartung* bevorzugen jedoch die Formalisierung mit Hilfe mathematischer Netzwerke oder Graphen, wie sie *Coupette* in ihrer Dissertation vorgedacht hatte. Im Netzwerk werden die »Objekte« zu »Knoten« und die Beziehungen zwischen »Objekten« zu Kanten.

Konkret wird über einen Rechtsstrukturvergleich zwischen Deutschland und den USA auf der Makroebene berichtet. Dabei geht es um die Frage: »Wird das nationale Gesetzes- und Verordnungsrecht in Deutschland und in den USA tatsächlich insgesamt immer umfangreicher, immer komplizierter und immer komplexer?« (S. 964). Die Antwort ist im Hinblick auf auf den Umfang relativ einfach. Da wird nur die Veränderung der Anzahl der Worte über einen Zeitraum von 20 Jahren gezählt. »Strukturelemente« ergeben sich aus der Gliederung er Rechtstexte, z. B. in Kapitel, Paragrafen oder Absätze. Referenzen sind Verweisungen und Bezugnahmen aller Art. Als wesentliches Ergebnis heben die Autoren hervor, dass in den USA die Verordnungen schneller gewachsen sind als die Gesetze, während s sich in Deutschland umgekehrt verhält. Vergleichsweise zeigt sich auch, dass in Deutschland innerhalb der Rechtstexte stärker mit Verweisungen gearbeitet wird, währen in den USA nicht selten Textwiederholungen anzutreffen sind. Das ist nur eine Andeutung. Im Detail zeigt sich ein genaueres und damit interessanteres Bild. Mir dient diese Arbeit als Beleg, dass ein Ähnlichkeitsvergleich, und zwar auch ein solcher mir Hilfe automatisierter Verfahren, möglich und aufschlussreich ist.

Zu 2: *Kilian Lüders/Bent Stohlmann, Classifying Proportionality*

Ziel der Arbeit von *Kilian Lüders* und *Bent Stohlmann* »Classifying Proportionality«[5] ist das *Argument Mining*. Es sollen domänenspezifische Argumentationstypen, im konkreten Fall das Verhältnismäßigkeitsargument, aus unstrukturierten Texten extrahiert werden. Es geht hier also um die von *Friedemann Vogel* so genannten »Mehrworteinheiten«, also um Textstellen, die wohl das gleiche Argument transportieren, dazu aber unterschiedliche Formulierungen verwenden, so dass sie mit einfachen Suchroutinen nicht aufgefunden werden können. Aber man kann auch nicht einfach irgendwelche KI mit einschlägigen Texten füttern, so dass sie selbsttätig Muster findet, die sich als die gesuchten Argumente erweisen. Vielmehr kann und muss KI für diese Aufgabe erst trainiert werden. Das ist die Lehre aus der Arbeit von *Lüders* und *Stohlmann.*

Die Autoren haben vergleichsweise verschiedene KI-Modelle eingesetzt, die mit strukturierten Texten trainiert werden. Ein wesentlicher Beitrag der Studie ist die

Erstellung und Veröffentlichung eines neuen Datensatzes. Dieser Datensatz besteht aus 300 Entscheidungen des BVerfG, in denen die auf Satzebene Verhältnismäßigkeitsargument annotiert wurden (insgesamt 54.929 Sätze). Mit diesem Datensatz wurden verschiedenen Modelle (u.a. neuronale Netze und Transformer wie BERT) trainiert, um vorherzusagen, ob ein Satz eine Verhältnismäßigkeitsprüfung beinhaltet. Trotz einer hohen Genauigkeit hatten die Modelle Schwierigkeiten, alle Sätze mit Verhältnismäßigkeitsargumenten zu identifizieren. Das beste Ergebnis erzielte das hybride BERT-BiLSTM-CRF-Modell. Das Ergebnis zeigt, dass die Identifizierung des Verhältnismäßigkeitsarguments komplex ist. Insbesondere in langen Passagen mit Verhältnismäßigkeitsprüfungen oder in Fällen, in denen das Argument nur angedeutet, aber nicht angewendet wird, treten Fehler auf.

Was die Quantifizierung des Verhältnismäßigkeitsarguments für die Rechtsdogmatik bringt, haben die Autoren in der Zeitschrift »Der Staat« ausgebreitet.[6]

Zu 3: Law Smells

Diese Arbeit finde ich besonders interessant, weil ein »struktureller« Vergleich zwischen Computer-Code und Rechtsnormen den Ausgangspunkt bildet. Vergleichen werden hier verschiedene Funktionen oder vielmehr Dysfunktionen bestimmter Gestaltungen von Computercode und Rechtsnormen. Es handelt sich wohl nicht mehr um einen bloßen Ähnlichkeitsvergleich, sondern um einen Funktionsvergleich. Das ist noch kein Beweis, aber immerhin in Hineis dafür, dass auch der Funktionsvergleich automatisiert werden könnte. Ein Beweis ist es deshalb nicht, weil mit dem *mapping* zunächst nur Äußerlichkeiten im Rechtscode gefunden werden. Auch die Suche nach Code Smells und Law Smells läuft nicht vollautomatisch. Dem Programm müssen die einschlägigen Muster vorgegeben werden.

Als Code Smell bezeichnen Softwareentwickler Programmteile, die zwar nicht unmittelbar zu Ausführungsfehlern führen, aber doch die Übersichtlichkeit und Pflege des Programms erschweren und damit die Fehlerwahrscheinlichkeit erhöhen. Das ist etwa der Fall, wenn ein Code oft wiederholt wird, »Klassen« zu viel leisten sollen und deshalb zu groß werden, Bausteine des Codes (»Methoden«) zu lang geraten, der Code mit Kommentaren versehen wird, weil Dritte ihn sonst schwer verstehen würden. Es gibt eine lange Liste gut definierter Code Smells, die bei einer Programmpflege (*refactoring*) möglichst beseitigt werden.

Coupette u. a. haben analog zu den Code Smells das Konzept von Law Smells entwickelt.[7] Law Smells sind oberflächliche Anzeichen, die auf tiefere Probleme hinweisen. Diese Probleme können es erschweren, die Norm zu verstehen und zu handhaben und sie bereiten Schwierigkeiten bei Rechtsänderungen. Die Autoren haben zehn solcher Probleme identifiziert, z. B.:

- Duplizierte Phrasen: Wiederholte Formulierungen, die zu Inkonsistenzen führen können.
- Lange Elemente: Sehr lange Abschnitte, die die Lesbarkeit und Strukturierung erschweren.
- Mehrdeutige Syntax: Verwendung von Operatoren oder Satzzeichen, die Raum für Interpretationen lassen.

Insoweit handelt es sich um statische Merkmale, das heißt solche, die aus dem Text allein erkennbar sind. Dynamisch sind dagegen das *term overloading* – derselbe Begriff wird an verschiedenen Stellen für unterschiedlich verwendet – oder eine Konstellation, die bei einer Reform *shotgun surgery*, das heißt Reparaturen an mehreren Stellen, erfordert. Die Autoren verwenden den United States Code (USC) als Beispieltext. Sie haben den USC nicht manuell annotiert, sondern setzen auf automatisierte Methoden zur Erkennung der vordefinierten von Law Smells. Hilfreich ist dabei die hierarchische Struktur des USC, bestehend aus Titeln, Kapiteln, Abschnitten und Unterabschnitten. So gelingt es, verschiedene Law Smells im USC zu identifizieren insbesondere komplexe Referenzstrukturen und lange Elemente wurden häufig gefunden. Die Arbeit von *Coupette* u. a. zu den Law Smells finde ich unter anderem deshalb interessant, weil die Autoren hier – ohne das besonders zu betonen – mit einer strukturellen Analogie arbeiten. Sie vergleichen Computercode mit Rechtsnormen und finden dabei einen ganzen Strauß gleichartiger Funktionen oder vielmehr Dysfunktionen. Damit öffnen sie ein Fenster für auch für die funktionale Rechtsvergleichung.

---

[1] *Corinna Coupette/Andreas M. Fleckner*, Quantitative Rechtswissenschaft, JZ 2018, 379–389.

[2] *Corinna Coupette/Janis Beckedorf/Dirk Hartung/Michael Bommarito/Daniel Martin Katz*, Measuring Law Over Time: A Network Analytical Framework with an Application to Statutes and Regulations in the United States and Germany, 2021, https://doi.org/10.48550/arXiv.2101.11284.

[3] Spaßeshalber habe ich Copilot gebeten, diesen Absatz zu prüfen. Copilot hat mir folgenden »Vorschlag für eine wissenschaftlich-publikationsfähige Fußnote« gemacht:

»Die von Coupette und Hartung betonte Komplexitätsreduktion wird als methodische Selbstverortung gegenüber kulturalistischen Einwänden präsentiert. Diese defensive Rahmung erscheint jedoch unnötig. Die Einsicht, dass Rechtssysteme komplexe adaptive Systeme darstellen, deren Gesamtzustand sich keinem Beobachter vollständig erschließt, gehört längst zum theoretischen Grundbestand der Rechts- und Sozialwissenschaften. Gerade weil wir in solchen Systemen handeln, sind wir epistemisch auf heuristische Ordnungsstrukturen angewiesen, die Komplexität zunächst grob bündeln und nur bei Bedarf vertiefen. Ein solcher Bedarf entsteht vor allem in Situationen individueller Entscheidungsrelevanz; doch selbst dann setzt jede vertiefte Analyse die Klärung übergeordneter struktureller Rahmenbedingungen voraus. Komplexitätsreduktion ist daher keine methodische Schwäche, sondern eine notwendige Voraussetzung systematischen Verstehens – und könnte entsprechend selbstbewusster vertreten werden.«

Weiter hat Copilot angeboten, die Fußnote »stärker theoretisch zu verankern (z. B. mit Luhmann, Simon, Gigerenzer, systems thinking)« und dazu folgendes Ergebnis geliefert.

»Die von Coupette und Hartung gewählte Rahmung ihrer Methode als Beitrag zur Komplexitätsreduktion wirkt unnötig defensiv. Seit Luhmanns grundlegender Einsicht, dass soziale Systeme ihre eigene Komplexität nur durch Reduktion operativ handhabbar machen können, gehört es zum theoretischen Gemeingut, dass vollständige Erfassung systemischer Dynamiken epistemisch ausgeschlossen ist (Luhmann, Soziale Systeme, 1987, insb. Kap. 3; ders., Ausdifferenzierung des Rechts, 1981). In komplexen adaptiven Systemen (im Sinne der Komplexitätsforschung) ist die Perspektivengebundenheit des Beobachters strukturell unvermeidlich; jede Analyse setzt daher eine selektive Fokussierung voraus (Miller/Page, Complex Adaptive Systems, 2007; De Domenico et al., Complexity Explained, 2019). Bereits Herbert Simon hat gezeigt, dass Akteure in komplexen Umwelten nur mittels »bounded rationality« und satisfizierender Heuristiken Orientierung gewinnen können (Herbert A. Simon, Models of Man, 1957; administrative Behavior, 4. Aufl. 1997). Gigerenzer hat diese Einsicht weiter zugespitzt: Gute Entscheidungen beruhen nicht auf vollständiger Informationsverarbeitung, sondern auf ökologisch rationalen Heuristiken, die Komplexität gezielt ausblenden (Gerd Gigerenzer, Adaptive Thinking, 2000; Rationality for Mortals, 2008). Auch im Systems Thinking gilt Komplexitätsreduktion nicht als epistemischer Mangel, sondern als notwendige Bedingung für Modellbildung und Intervention (Meadows, Thinking in Systems, 2008). Vor diesem Hintergrund ist Komplexitätsreduktion nicht bloß ein methodisches Zugeständnis, sondern eine epistemische Notwendigkeit jeder rechtsvergleichenden Analyse. Eine Methode, die dies offen anerkennt, könnte dies mit größerer Selbstverständlichkeit und weniger apologetischem Unterton tun.«

[4] *Jakob Freund/Bernd Rücker*, Praxishandbuch BPMN, 7. Aufl. 2025.

[5] *Kilian Lüders/Bent Stohlmann*, Classifying Proportionality – Identification of a Legal Argument, Artificial Intelligence and Law 2025, 1051–1078.

[6] Bent Stohlmann/Kilian Lüders/Luisa Wendel/Leonard Hoeft/Sophie Reule/Alexander Tischbirek, Konsolidierung statt Siegeszug. Eine quantitative Untersuchung der Verwendung des Grundsatzes der Verhältnismäßigkeit in der Rechtsprechung des Bundesverfassungsgerichts, Der Staat 2024, 217–252.

[7] *Corinna Coupette/Dirk Hartung/Janis Beckedorf/Maximilian Böther/Daniel Martin Katz*, Law Smells, Artificial Intelligence Law 31, 2023, 335–368.

Ähnliche Themen

- Rechtsvergleich mit Hilfe Künstlicher Intelligenz (KI)
- Vergleichsobjekte und Vergleichsmengen
- Der Vergleich des Vergleichs als Weg zur Interdisziplinarität
- Frege und die Frage nach der Intelligenz der künstlichen Intelligenz
- Moralische Maschinen 2.0
- Recht muss anthropozentrisch bleiben – oder soll es Menschen künstlicher Intelligenz ausliefern?
- Digitalisierung der Rechtskommunikation – der Fortschritt ist eine Schnecke
- Travelling Models VIII: Nun kommt man auch in Frankfurt auf den Trichter.
- Travelling Models I: Rechtsvergleichung